

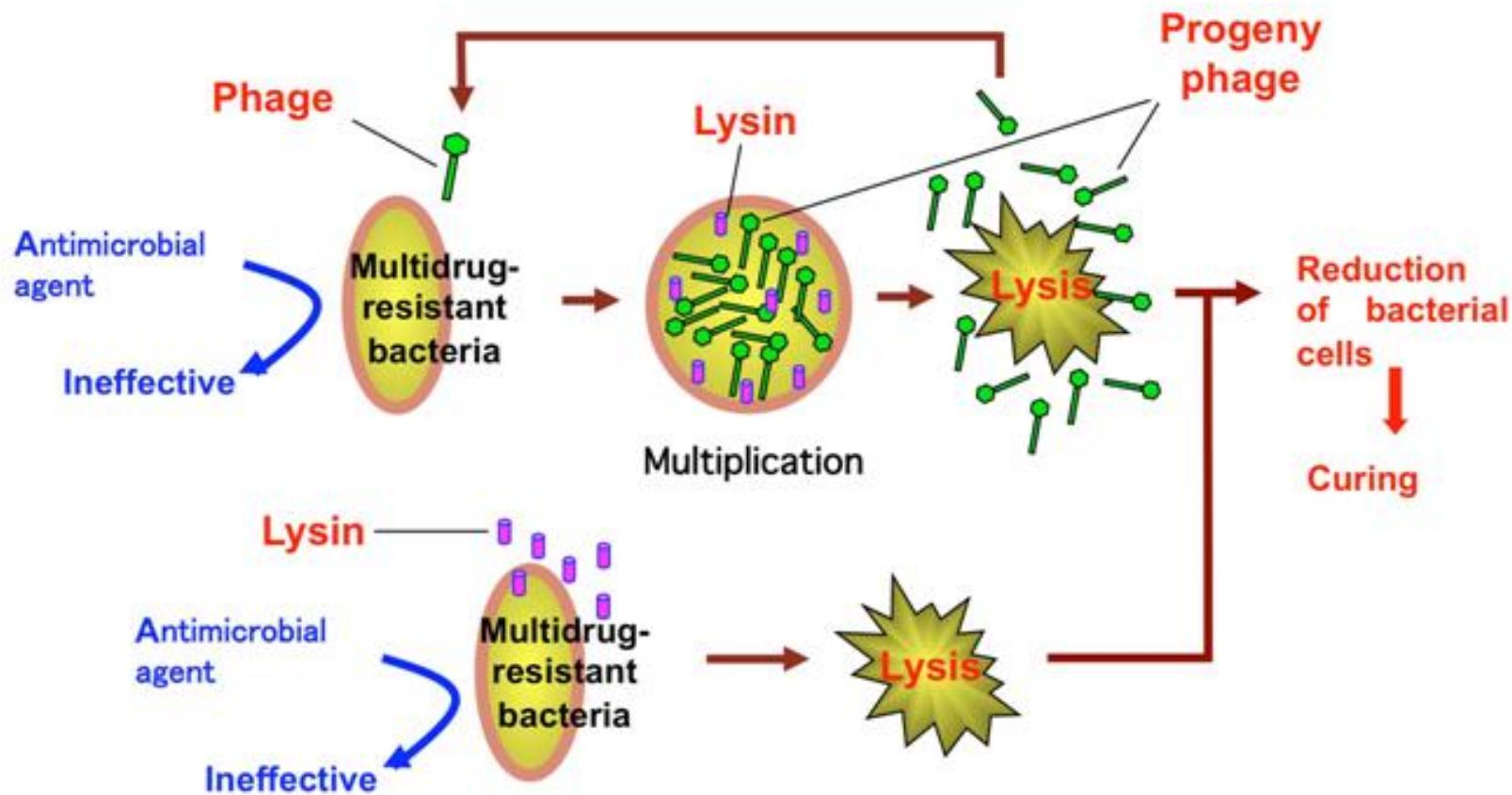
# The Rumen Virome. Genomic Approaches for Investigating Evolution of Viral Populations of the Rumen and Implications for Antibiotic Resistance

*Bryan A. White*

*Professor, Institute for Genomic Biology  
Director, Mayo Clinic/University of Illinois  
Strategic Alliance for Technology-Based Healthcare  
Department of Animal Sciences  
University of Illinois at Urbana-Champaign*

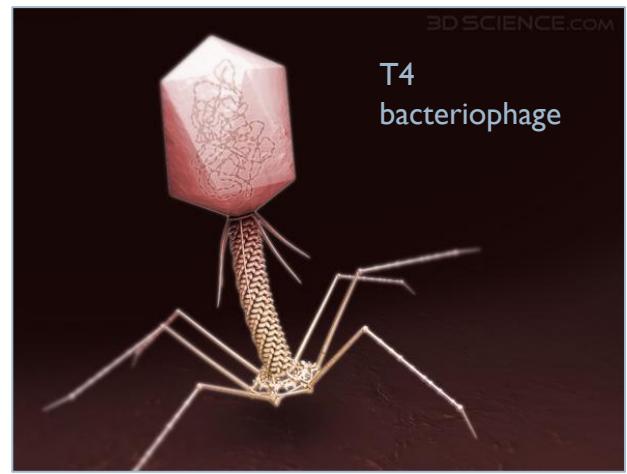


# Phage Therapy



# Viruses

- ▶ Most abundant biological entities on the planet
  - ▶  $\sim 10^{31}$  total viral particles
  - ▶ 5-10 viral particles per bacterial cell
- ▶ Largely uncharacterized
  - ▶ “natural viral communities probably represent the largest unexplored area of genetic information space left on the planet” –Rohwer et al., 2009



(3dscience.com)

# Role of Viruses in Evolution

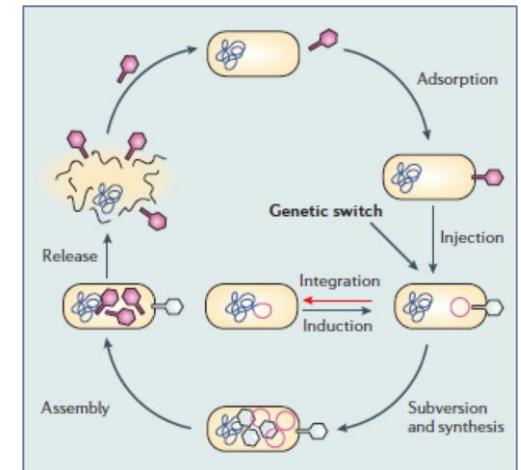
---

- ▶ **Control microbial populations**
  - ▶ Maintain microbial populations at levels that do not overwhelm the ecosystem
  - ▶ Maintain microbial diversity
  - ▶ Drive evolution through natural selection of phage-resistant microbes



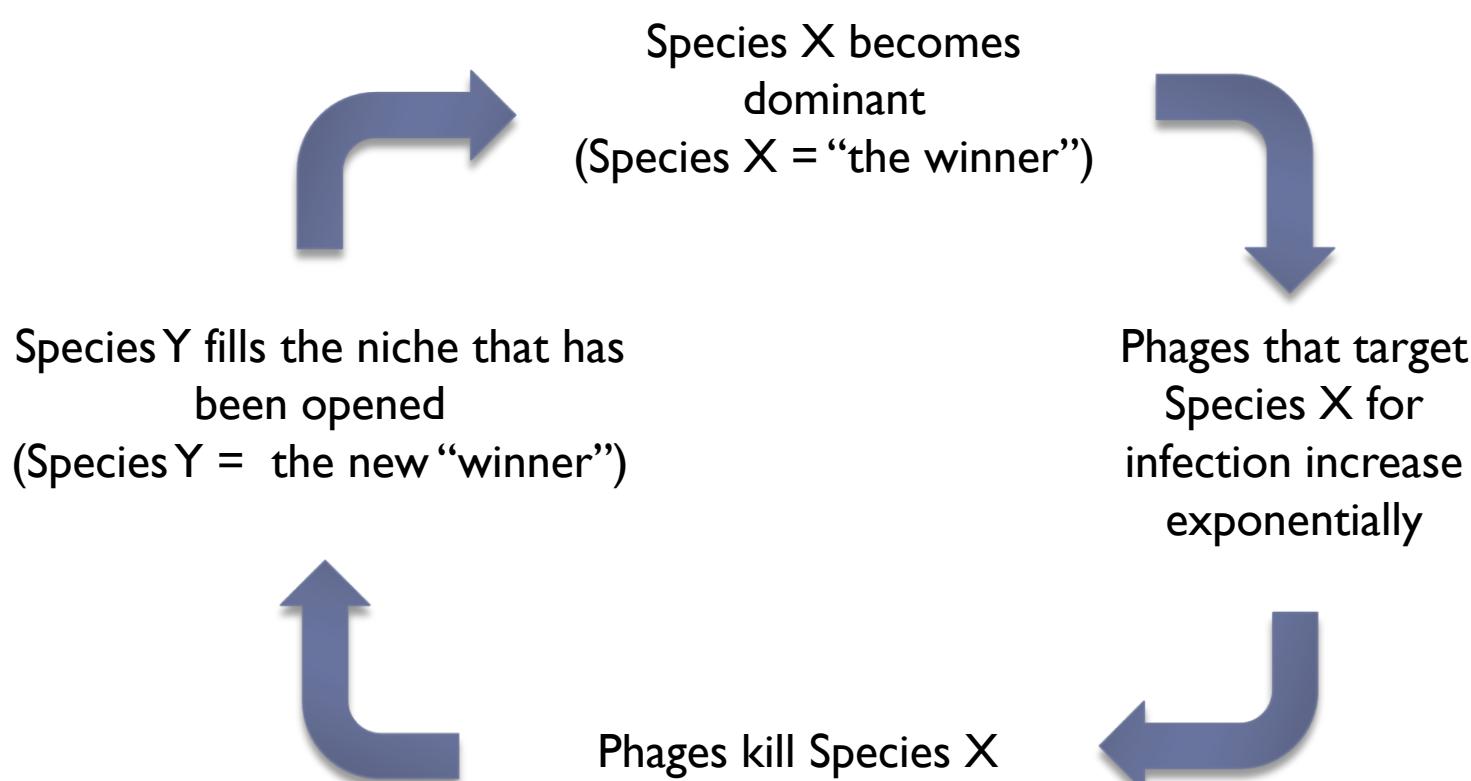
# Role of Viruses in Evolution

- ▶ Control microbial populations
  - ▶ Maintain microbial populations at levels that do not overwhelm the ecosystem
  - ▶ Maintain microbial diversity at ecological scales
  - ▶ Drive evolution through natural selection of phage-resistant microbes
  
- ▶ Facilitate horizontal gene transfer
  - ▶ Primarily through gain and loss of prophages



(Sturino et al., 2006)

# “Kill-the-Winner” Dynamics



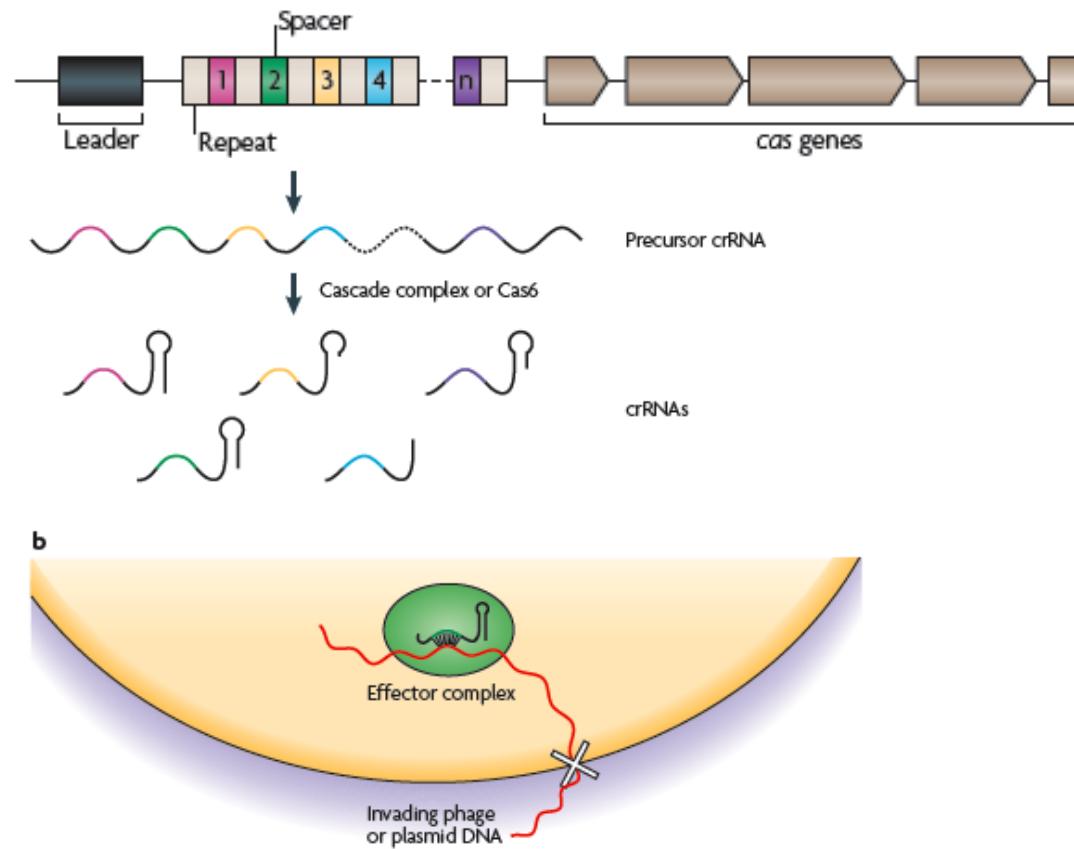
Conceptual and mathematical models suggest that viruses can influence the winner in competitions for nutrients. In this scenario, once a microbe becomes dominant, the increase in abundance increases its contacts with viruses leading to significant increases in infection and subsequent lysis, which then control its abundance. As a consequence, viruses may sustain a high population diversity amongst their hosts by controlling the growth of organisms that would extensively proliferate due to advantages in nutrient acquisition. This would allow for the survival of less competitive but virus-resistant microbes.



Rodriguez-Valera et al., 2009

# The CRISPR/cas System

- ▶ Acquired “immune” system for phage resistance
  - ▶ Clustered Regularly Interspaced Short Palindromic Repeats

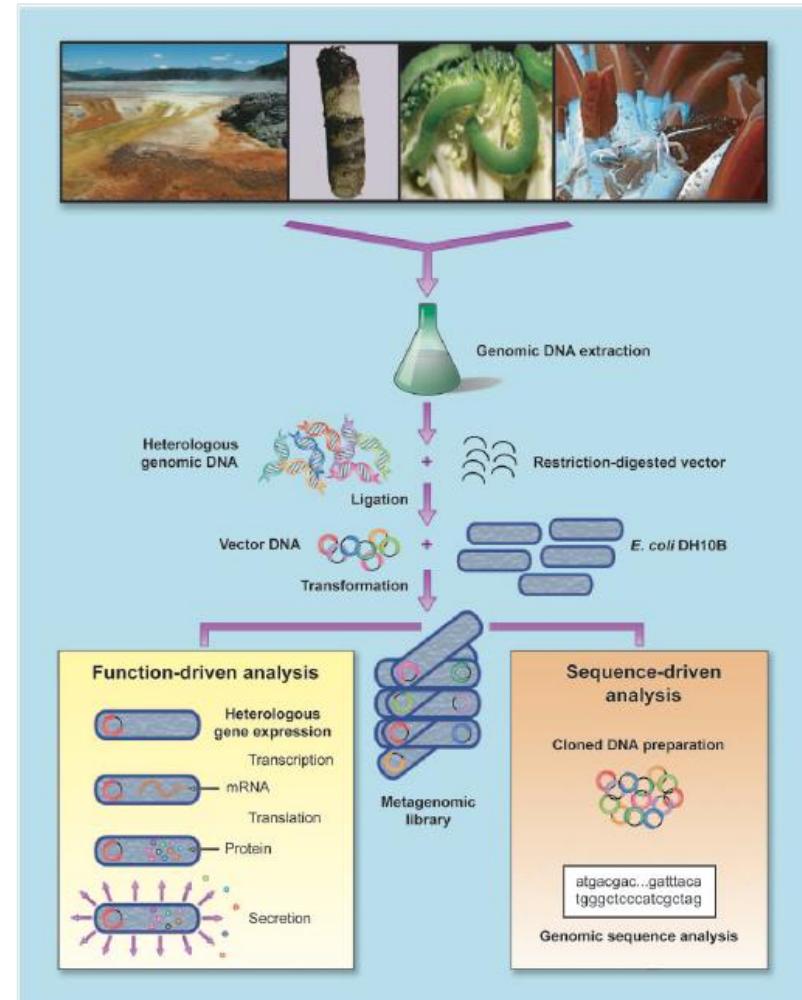


Marraffini and Sontheimer, 2010

# Metagenomics

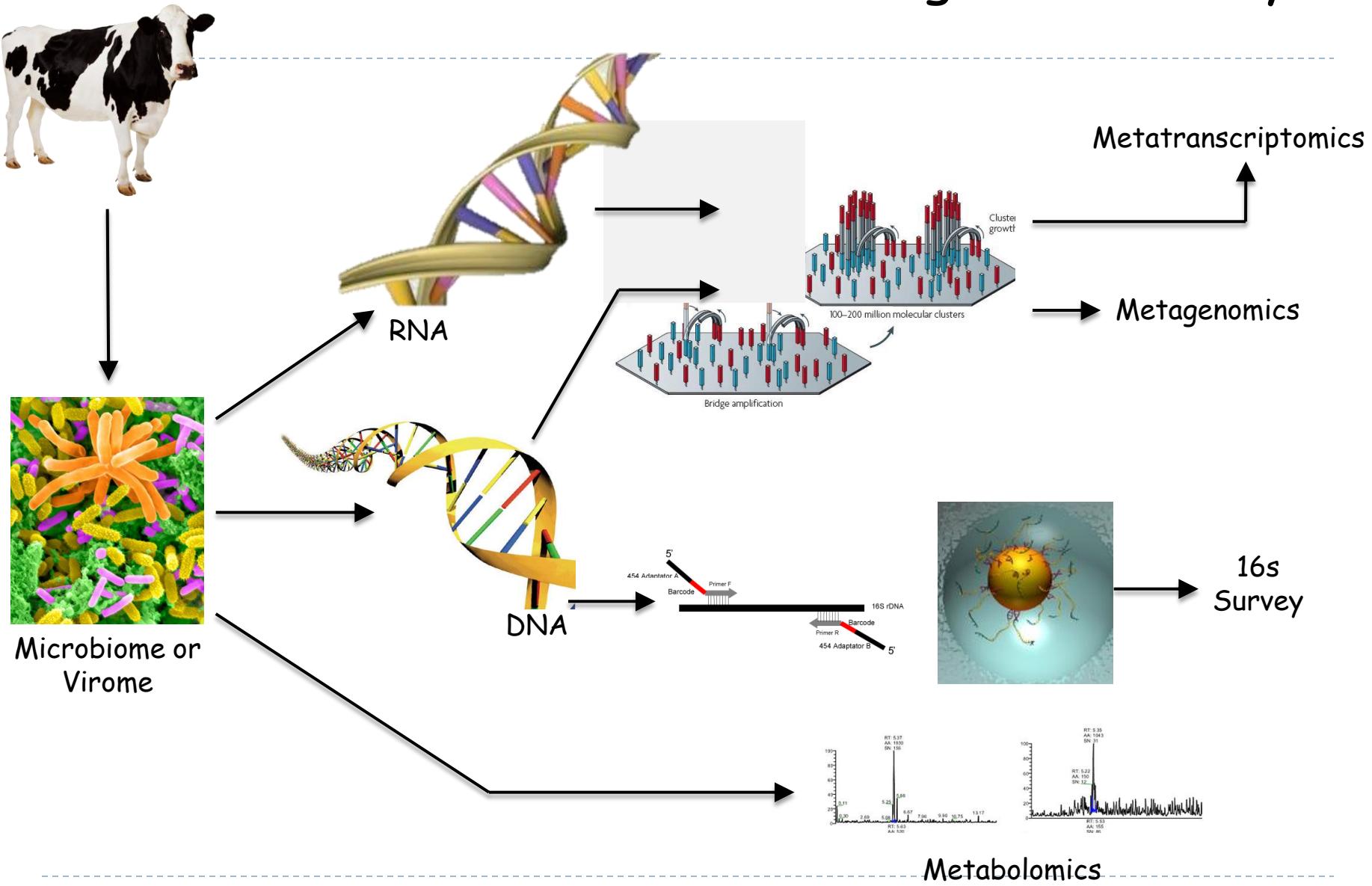
- ▶ The functional and sequence-based analysis of the collective genomes contained in an environmental sample  
-Riesenfeld et al., 2004

- ▶ Who's there?
- ▶ What are they doing?
- ▶ Culture independent

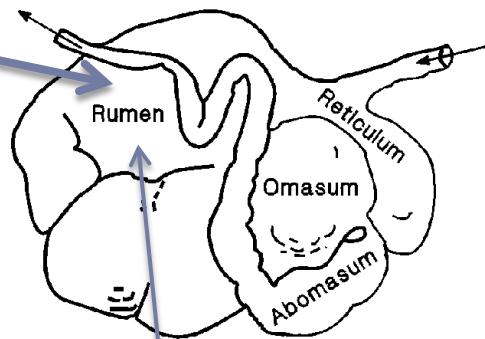
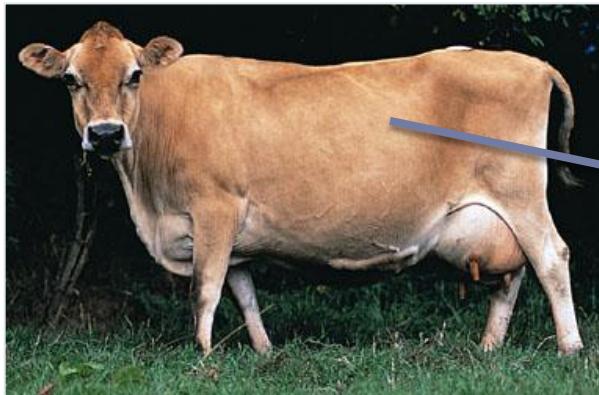


(J. Handelsman, 2004)

# Microbiome and Metagenomic Analysis

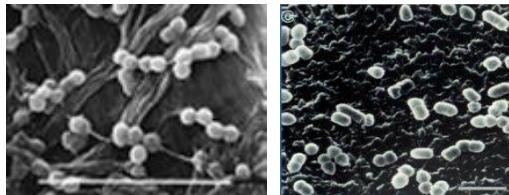


# The Rumen Ecosystem

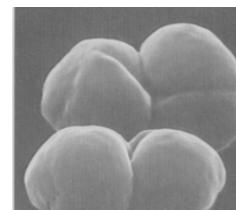


Bacteriophages –  
 $10^7$ - $10^9$  particles/ml

Bacteria  
 $10^{10}$ - $10^{11}$  cells/g



Archaea



Eukarya



Fungi

Protozoa



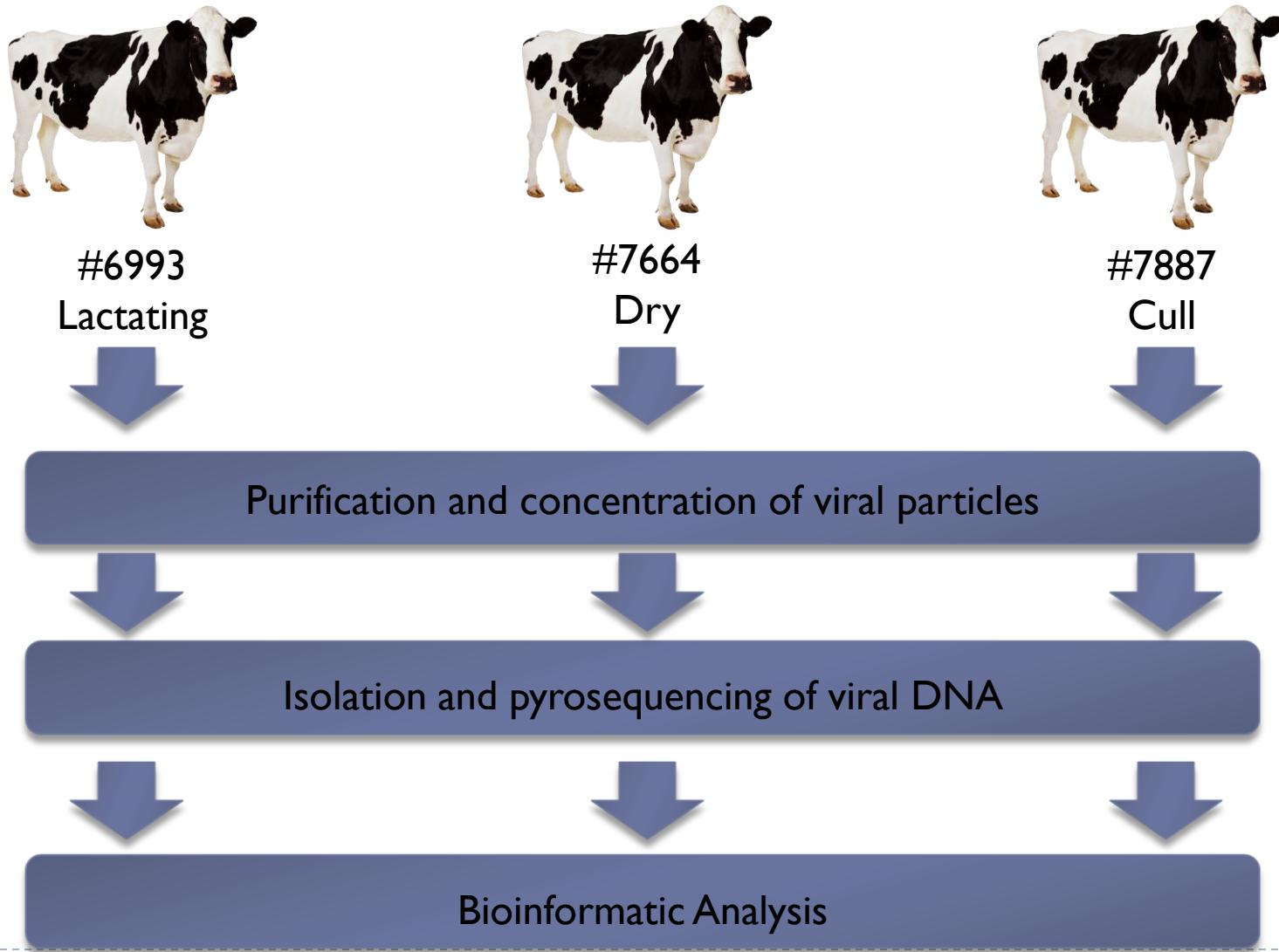
# Horizontal Gene Transfer (HGT) in the Rumen

---

- ▶ Antibiotic resistance genes
  - ▶ McCuddin *et al.*, 2006
  - ▶ Toomey *et al.*, 2009
- ▶ Transfer of genes for survival/adaptation between bacteria and rumen Ciliates
  - ▶ Ricard *et al.*, 2006
- ▶ Natural transformation in *Ruminococcus* spp.
  - ▶ Klieve *et al.*, 2005
- ▶ Bacteriophages play an integral role in the community dynamics of the rumen and drive the evolution of the rumen microbiome by means of phage-mediated horizontal gene transfer.



# Methods



# Summary of Pyrosequence Data

	Cow i.d.		
	<u>6993</u>	<u>7664</u>	<u>7887</u>
Production stage	Lactating	Dry	Cull
Total sequence reads	329,384	540,960	298,062
Total sequence size (bp)	129,599,968	196,184,178	99,437,869
Average read length (bp)	394	363	334
Shortest read length (bp)	100	100	100
Longest read length (bp)	673	615	615



# PHACCS Diversity Estimates

	Cow i.d.		
	<u>6993</u>	<u>7664</u>	<u>7887</u>
Best Model <sup>a</sup>	Logarithmic	Power	Lognormal
Error <sup>b</sup>	0.675	0.277	0.702
Richness	4080 genotypes	22100 genotypes	27800 genotypes
Evenness	0.988	0.945	0.834
Most Abundant Genotype	0.949% of the community	0.829% of the community	1.53% of the community
Shannon-Wiener Index	8.22	9.46	8.53

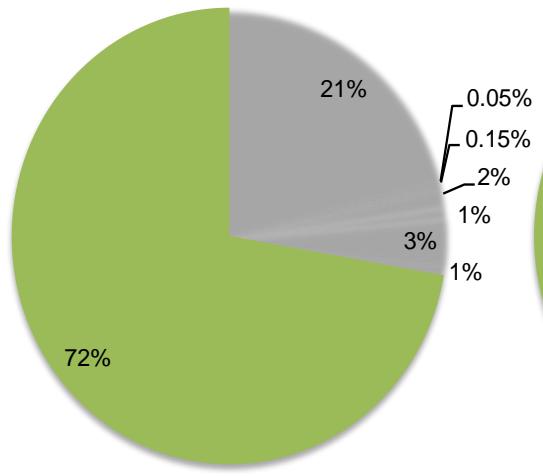
<sup>a</sup>The best model is the one with the smallest error.

<sup>b</sup>The error is the quantification of the difference between the predicted contig spectrum and the experimental contig spectrum.

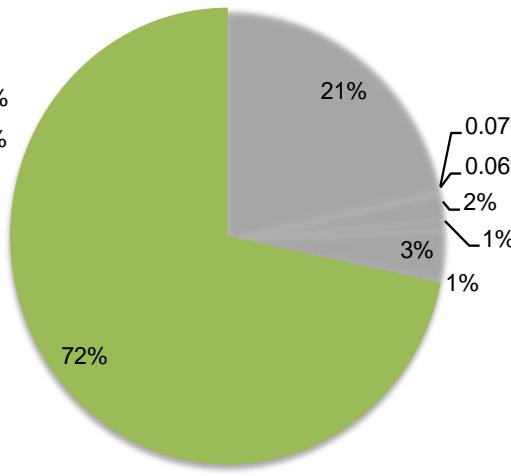


# Virome Comparison to NR Database

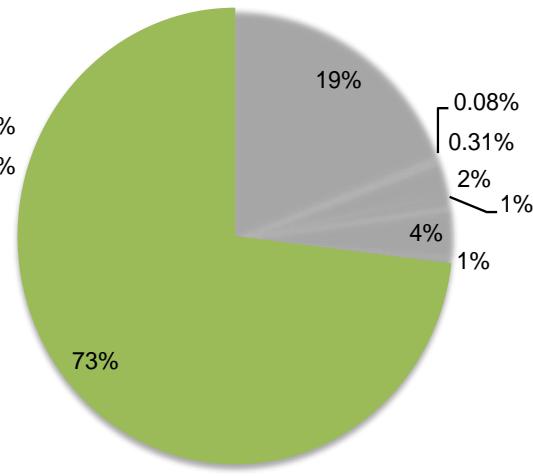
Cow 7887



Cow 7664



Cow 6993



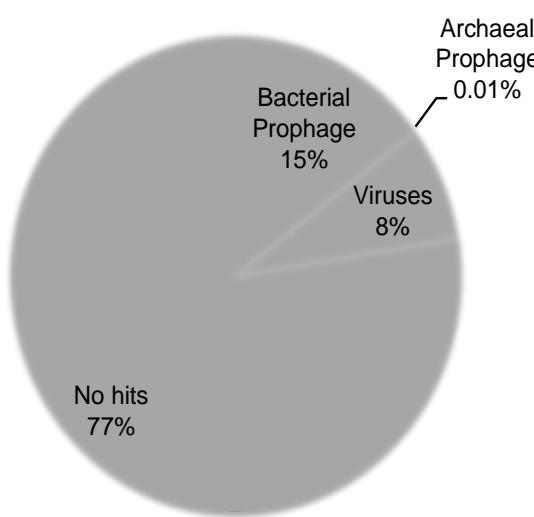
- Bacteria
- Eukaryota
- Cellular organisms
- Not assigned
- Archaea
- Viruses
- Root
- No hits

BLASTx, E-value < 0.001

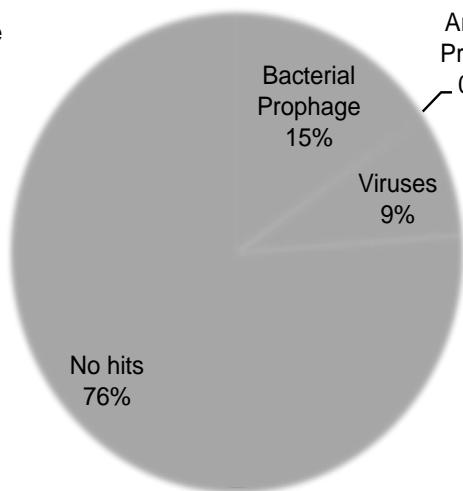


# Virome Comparison to Viral Database

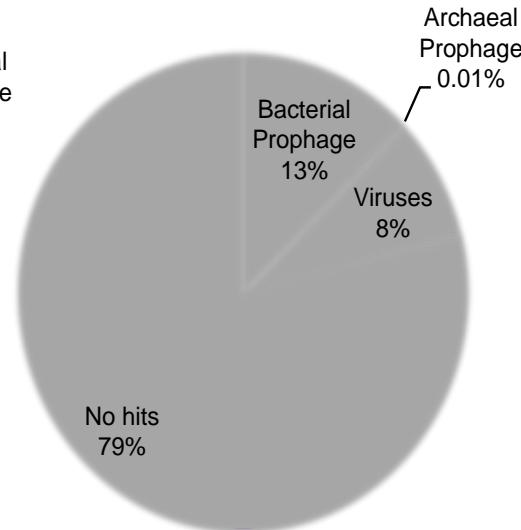
**Cow 7887**



**Cow 7664**



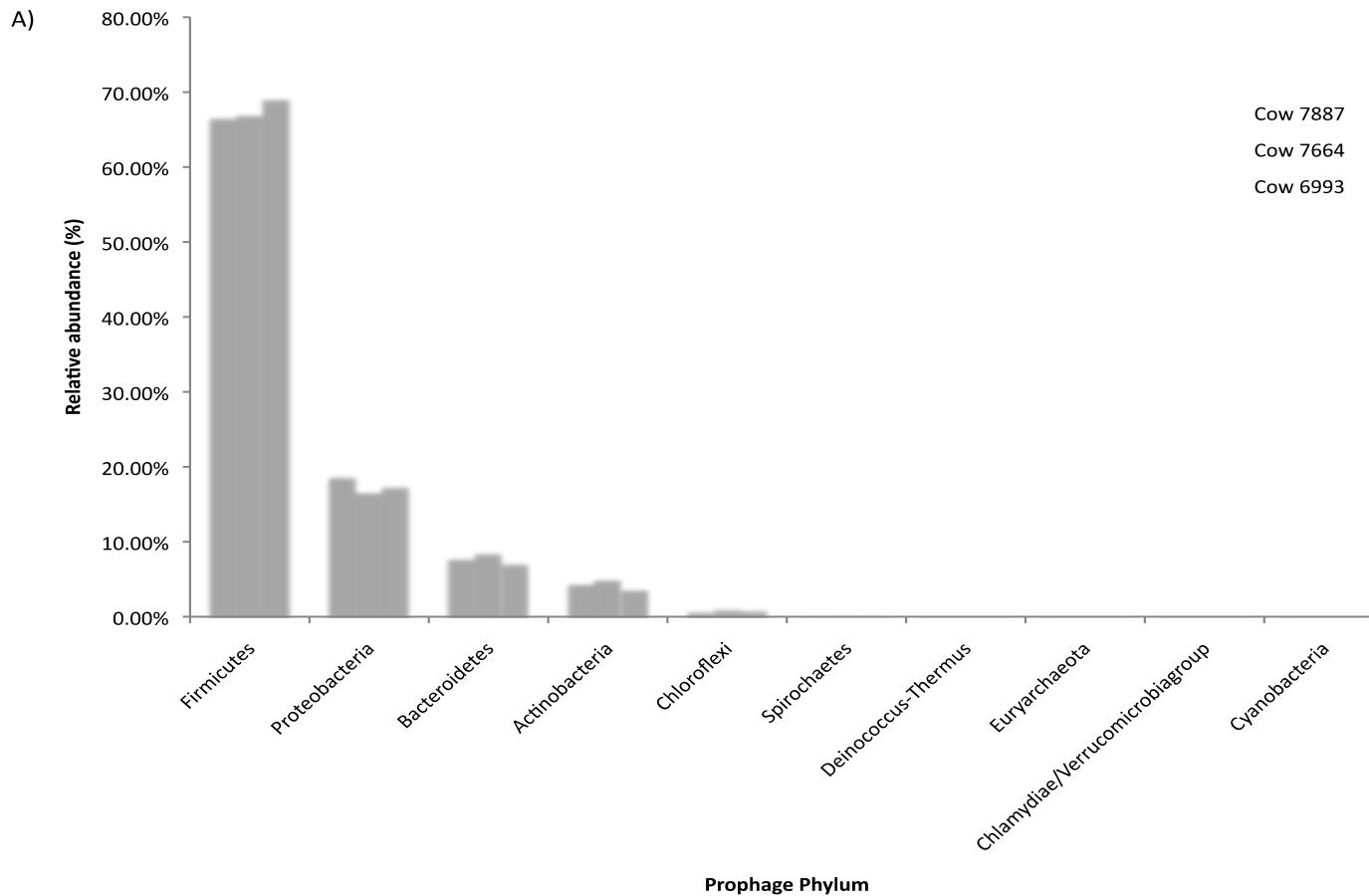
**Cow 6993**



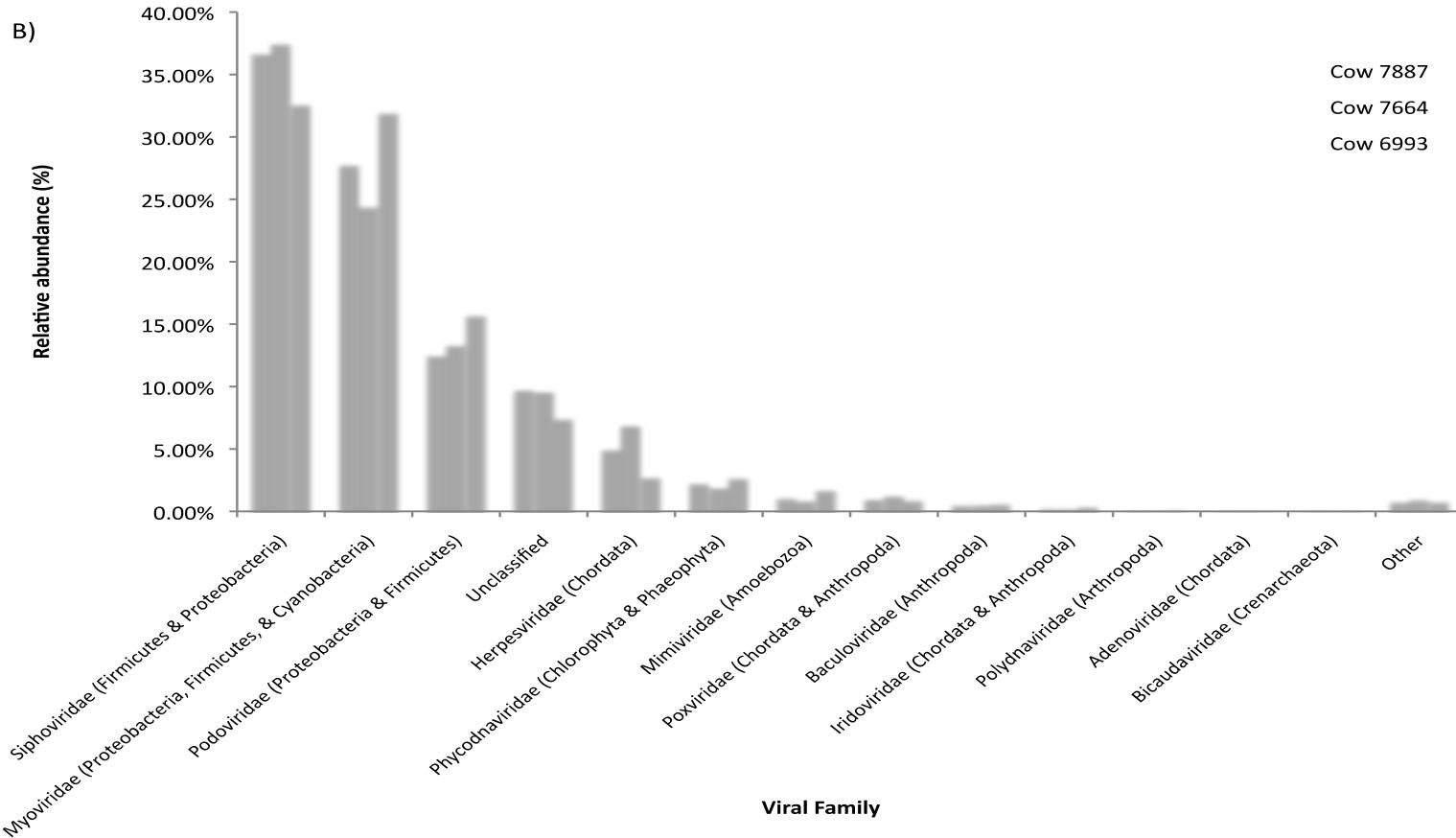
tBLASTx, E-value < 0.001



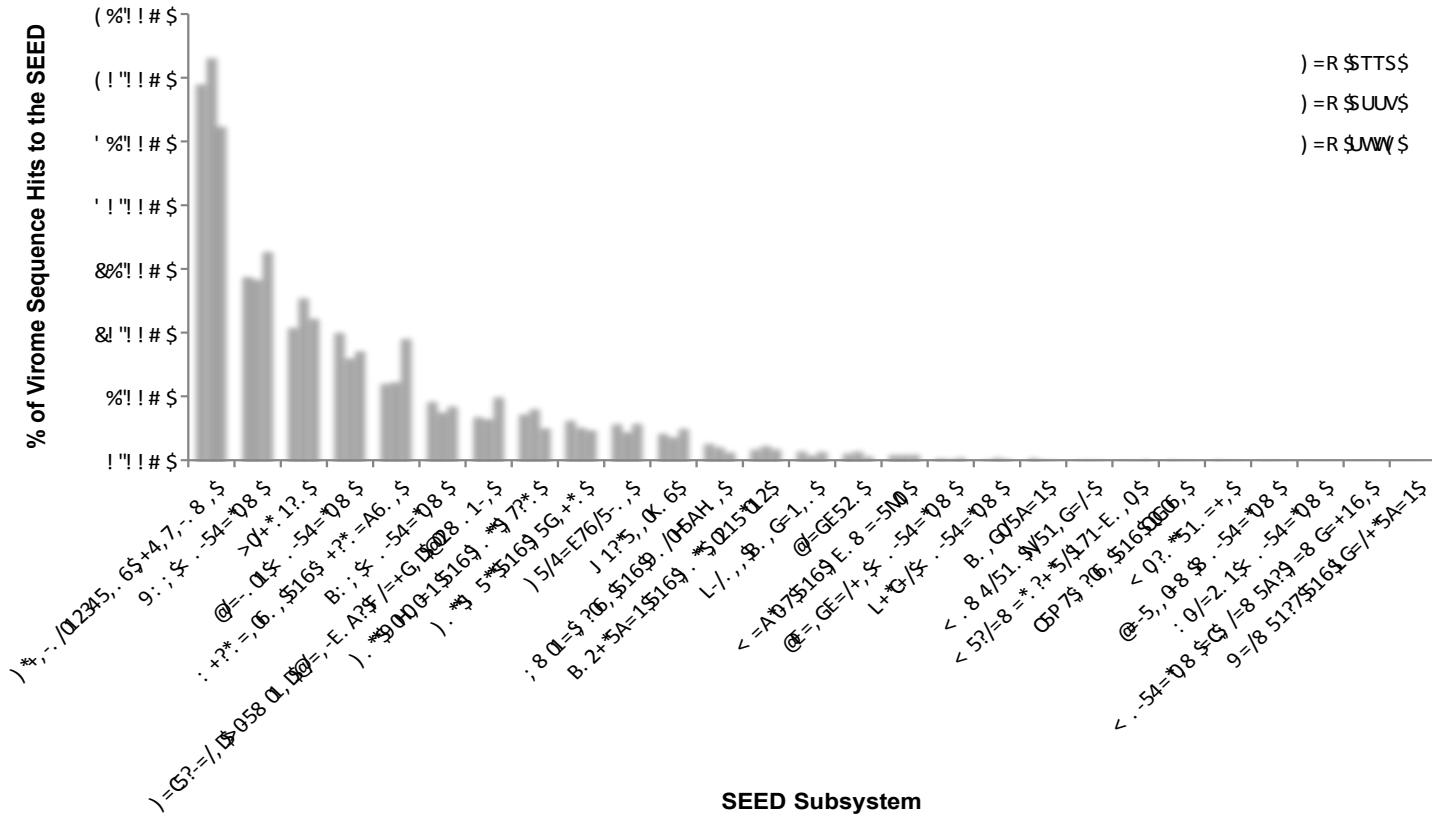
# Distribution of Prophage Phyla



# Distribution of Viral Families



# SEED Subsystems



E-value < 10<sup>-5</sup>

# Summary

---

- ▶ Rumen viral population is abundant and diverse
- ▶ Bacteriophage populations are infecting the predominant bacterial phyla (**Firmicutes, Proteobacteria, Bacteroidetes**)
  - ▶ Maintenance of a balanced microbial community
- ▶ Phages are carrying genes for DNA metabolism, virulence, and protein metabolism
  - ▶ Low abundance of Carbohydrates and Amino Acids protein families compared to study by Dinsdale *et al.*, 2008
- ▶ Phages could represent a shared gene pool for the rumen ecosystem
  - ▶ Prophages were abundant





# INTERACTIONS BETWEEN THE RUMEN VIRAL METAGENOME AND THE RUMEN MICROBIOME

# Phage–bacteria relationships and CRISPR elements revealed by a metagenomic survey of the rumen microbiome

Margret E. Berg Miller,<sup>1</sup> Carl J. Yeoman,<sup>1</sup>  
Nicholas Chia,<sup>1,2</sup> Susannah G. Tringe,<sup>3</sup>  
Florent E. Angly,<sup>4</sup> Robert A. Edwards,<sup>5</sup>  
Harry J. Flint,<sup>6</sup> Raphael Lamed,<sup>7</sup> Edward A. Bayer<sup>8</sup>  
and Bryan A. White<sup>1\*</sup>

<sup>1</sup>Institute for Genomic Biology, University of Illinois at Urbana-Champaign, 1206 W. Gregory Dr, Urbana, IL 61801, USA.

<sup>2</sup>Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green St., Urbana, IL 61801, USA.

<sup>3</sup>DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA.

<sup>4</sup>Australian Centre for Ecogenomics, University of Queensland, St. Lucia, Brisbane, Australia.

<sup>5</sup>Department of Computer Science, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA.

<sup>6</sup>Microbial Ecology Group, Rowett Institute of Nutrition and Health, University of Aberdeen, Aberdeen, UK.

<sup>7</sup>Department of Molecular Microbiology and Biotechnology, Tel Aviv University, Ramat Aviv, Israel.

<sup>8</sup>Department of Biological Chemistry, The Weizmann Institute of Science, Rehovot, Israel.

not match any previously described virus. Prophages outnumbered lytic phages approximately 2:1 with the most abundant bacteriophage and prophage types being associated with members of the dominant rumen phyla (*Firmicutes* and *Proteobacteria*). Metabolic profiling based on SEED subsystems revealed an enrichment of sequences with putative functional roles in DNA and protein metabolism, but a surprisingly low proportion of sequences assigned to carbohydrate and amino acid metabolism. We expanded our analysis to include previously described metagenomic data and 14 reference genomes. Clustered regularly interspaced short palindromic repeats (CRISPR) were detected in most of the microbial genomes, suggesting previous interactions between viral and microbial communities.

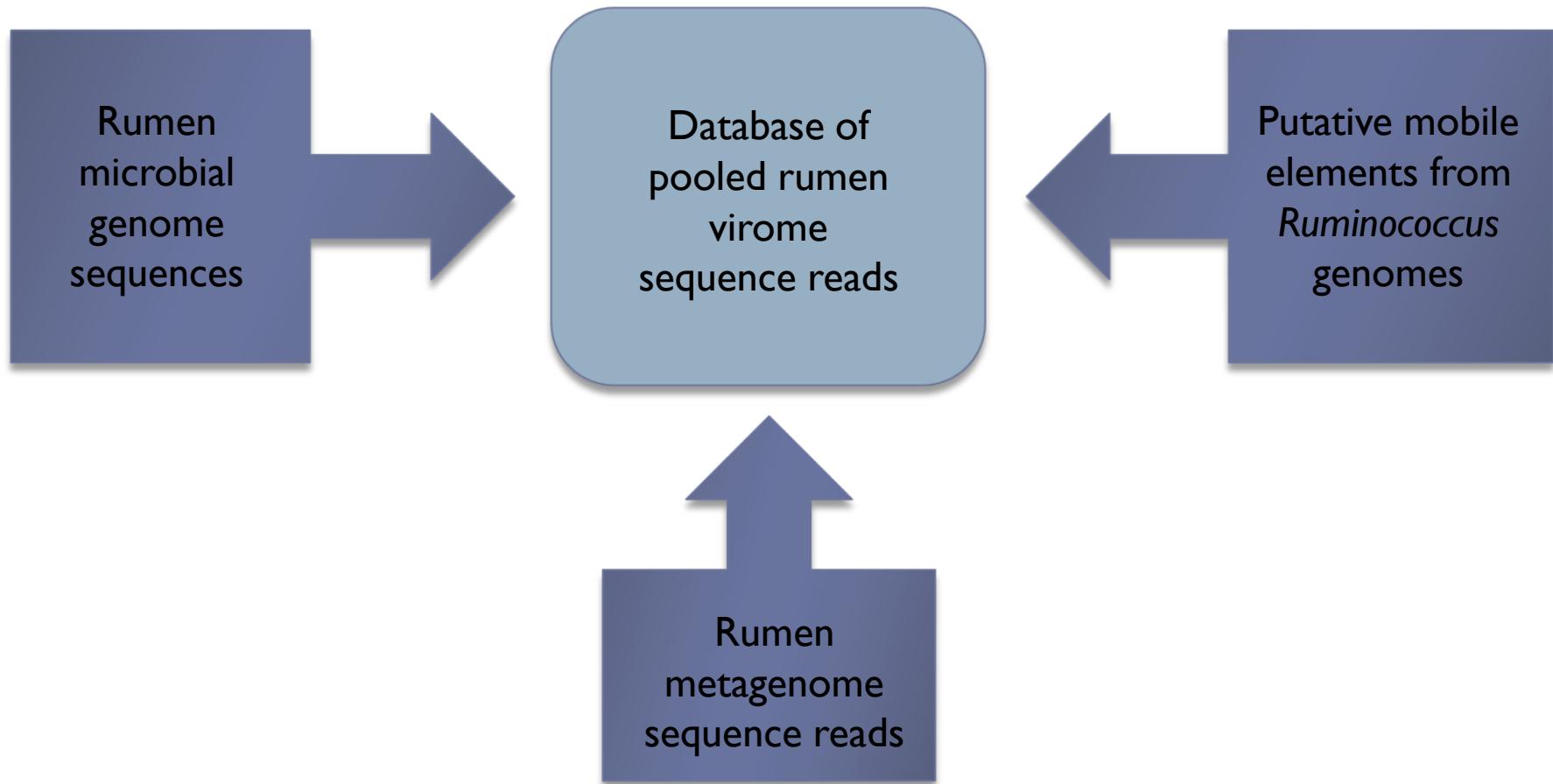
## Introduction

The rumen environment is home to a diverse population of microbes encompassing all three domains of life: *Bacteria*, *Archaea* and *Eukarya*. Of the three domains of life inhabiting the rumen, the bacteria are predominant ( $10^{10}$ – $10^{11}$  cells per gram of rumen contents) (Hespell *et al.*, 1997; Russell and Rychlik, 2001). Previous research has

# Methods

---

tBLASTx, E-value < 0.001



**Table 3.** Microbial reference genomes used in this study.

Species <sup>a</sup>	Status	Genome size (Mb)	Availability	Reference
<i>Actinobacillus succinogenes</i> 130Z	Closed	2.32	CP000746 <sup>b</sup>	
<i>Butyrivibrio proteoclasticus</i> B316:				
Chromosome 1	Closed	3.55	CP001810 <sup>b</sup>	Kelly <i>et al.</i> (2010)
Chromosome 2	Closed	0.3	CP001811 <sup>b</sup>	Kelly <i>et al.</i> (2010)
pCY186	Closed	0.19	CP001813 <sup>b</sup>	Kelly <i>et al.</i> (2010)
pCY360	Closed	0.36	CP001812 <sup>b</sup>	Kelly <i>et al.</i> (2010)
<i>Fibrobacter succinogenes</i> S85	Closed	3.84	FibRumBa <sup>c</sup>	
<i>Mannheimia succiniciproducens</i> MBEL55E	Closed	2.31	AE016827 <sup>b</sup>	Hong <i>et al.</i> (2004)
<i>Methanobrevibacter ruminantium</i> M1	Closed	2.9	CP001719 <sup>b</sup>	Leahy <i>et al.</i> (2010)
<i>Prevotella bryantii</i> B <sub>1</sub> 4	Draft	3.8	ADWO01000000 <sup>b</sup>	Purushe <i>et al.</i> (2010)
<i>Prevotella ruminicola</i> 23	Closed	3.62	CP002006 <sup>b</sup>	Purushe <i>et al.</i> (2010)
<i>Ruminococcus albus</i> 7:				
Chromosome	Closed	3.68	CP002403 <sup>b</sup>	
pRUMAL01	Closed	0.42	CP002404 <sup>b</sup>	
pRUMAL02	Closed	0.35	CP002405 <sup>b</sup>	
pRUMAL03	Closed	0.016	CP002406 <sup>b</sup>	
pRUMAL04	Closed	0.007	CP002407 <sup>b</sup>	
<i>Ruminococcus albus</i> 8	Draft	4.2	FibRumBa <sup>c</sup>	
<i>Ruminococcus flavefaciens</i> 007C	Draft	3.64	Sanger Institute <sup>d</sup>	
<i>Ruminococcus flavefaciens</i> 17	Draft	3.45	AFNE00000000	This study
<i>Ruminococcus flavefaciens</i> FD-1	Draft	4.57	ACOKoooooooooooo	Berg Miller <i>et al.</i> , 2009
<i>Ruminococcus</i> sp. 18P13	Draft	2.57	FP929052 <sup>b</sup>	
<i>Wolinella succinogenes</i> DSM 1740	Closed	2.11	BX571656 <sup>b</sup>	Baar <i>et al.</i> (2003)

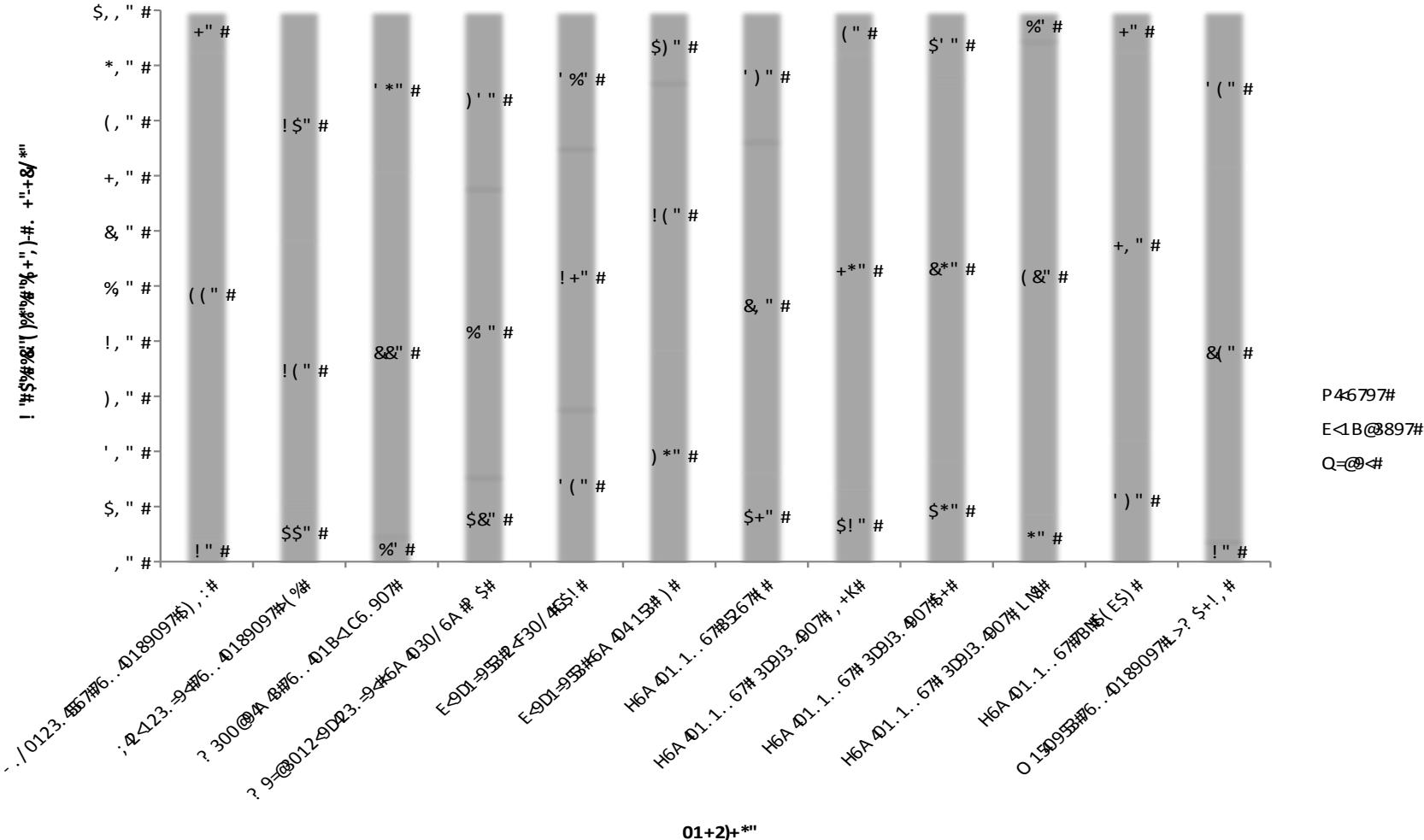
a. All species originated from the bovine rumen, except *Ruminococcus* sp. 18P13, which is a human intestinal isolate.

b. Available from GenBank.

c. FibRumBa database (<http://jcvb.org/rumenomics>).

d. Wellcome Trust Sanger Institute (<http://www.sanger.ac.uk>).

# Microbial Genomes vs. Virome

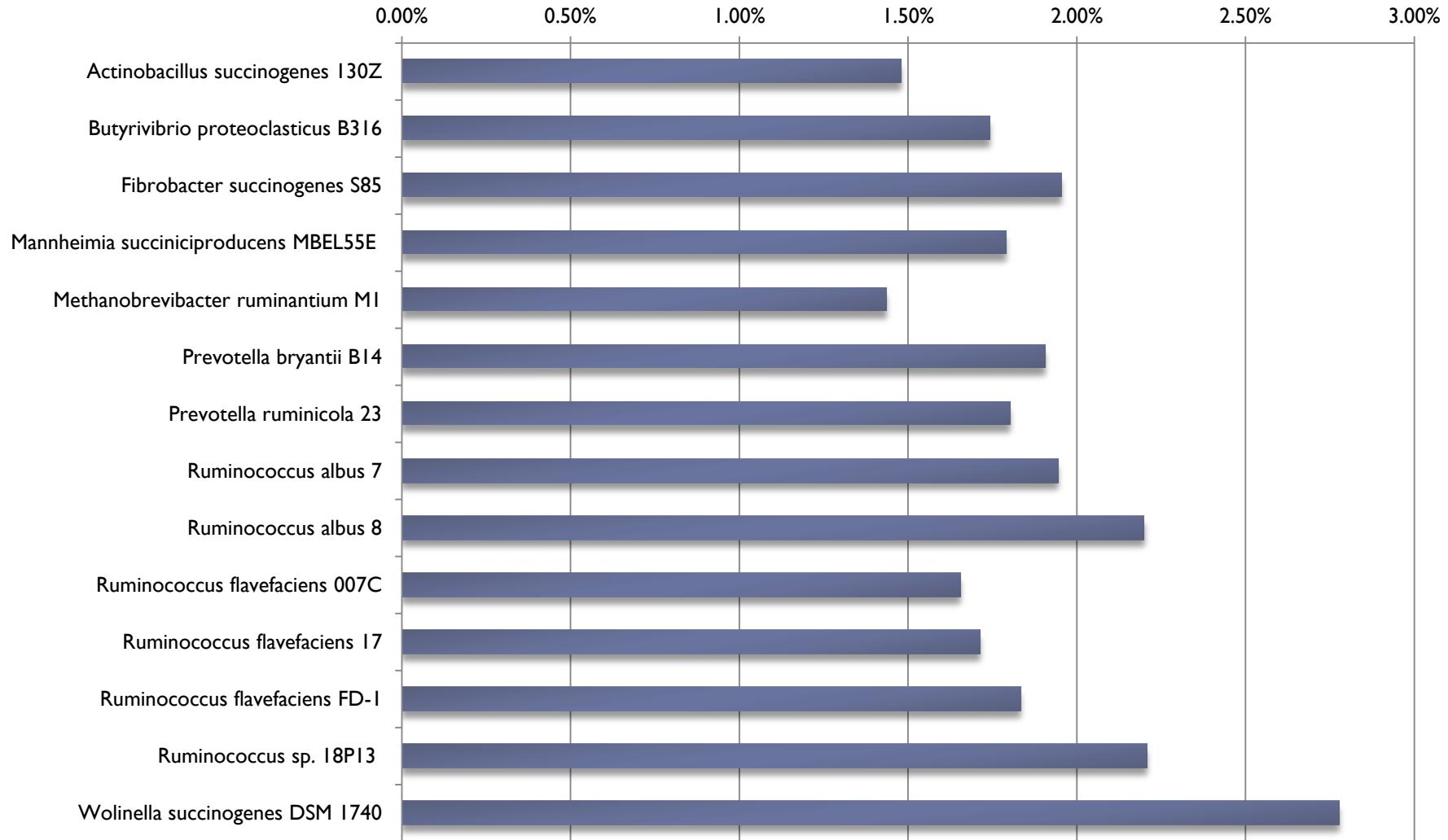


# Mobile Elements Detected in Rumen Microbial Genomes

Species	Total Elements	Phage-related	Integrase	Recombinases	Transposable Elements	Other
<i>Actinobacillus succinogenes</i> 130Z	26	18 (69)	2 (8)	2 (8)	2 (8)	3 (12)
<i>Butyrivibrio proteoclasticus</i> B316	47	6 (13)	6 (13)	5 (11)	33 (70)	1 (2)
<i>Fibrobacter succinogenes</i> S85	7	2 (29)	0 (0)	2 (29)	2 (29)	1 (14)
<i>Mannheimia succiniciproducens</i> MBEL55E	28	18 (64)	0 (0)	2 (7)	7 (25)	2 (7)
<i>Methanobrevibacter ruminantium</i> M1	17	6 (35)	2 (12)	1 (6)	9 (53)	0 (0)
<i>Prevotella bryantii</i> B14	34	0 (0)	12 (35)	4 (12)	16 (47)	3 (9)
<i>Prevotella ruminicola</i> 23	17	8 (47)	5 (29)	1 (6)	1 (6)	3 (18)
<i>Ruminococcus albus</i> 7	132	22 (17)	31 (23)	35 (27)	64 (48)	7 (5)
<i>Ruminococcus albus</i> 8	91	10 (11)	12 (13)	19 (21)	52 (57)	3 (3)
<i>Ruminococcus flavefaciens</i> 007C	49	26 (53)	12 (24)	15 (31)	2 (4)	5 (10)
<i>Ruminococcus flavefaciens</i> 17	39	18 (46)	9 (23)	12 (31)	9 (23)	3 (8)
<i>Ruminococcus flavefaciens</i> FD-I	109	49 (45)	22 (20)	33 (30)	13 (12)	13 (12)
<i>Ruminococcus</i> sp. 18P13 (human isolate)	31	17 (55)	7 (23)	4 (13)	3 (10)	3 (10)
<i>Wolinella succinogenes</i> DSM 1740	20	2 (10)	2 (10)	1 (5)	15 (75)	2 (10)



# Percent Abundance of Features in the Resistance to Antibiotics and Toxic Compounds Subsystem



**Table 5.** Microbial genomes and genome bins used in this study and the number of CRISPR loci and Cas proteins detected.

Species/genome bin	Total predicted CRISPR loci <sup>a</sup>	CRISPR-like structures <sup>a</sup>	Questionable CRISPR structures <sup>a</sup>	Sets of Cas proteins <sup>b</sup>
<i>Actinobacillus succinogenes</i> 130Z	2	2	0	2
<i>Butyrivibrio proteoclasticus</i> B316	8	1	7	0
Chromosome 1	7	0	7	0
Chromosome 2	0	0	0	0
pCY186	0	0	0	0
pCY360	1	1	0	0
<i>Fibrobacter succinogenes</i> S85	15	6	9	2
<i>Mannheimia succiniciproducens</i> MBEL55E	7	3	4	2
<i>Methanobrevibacter ruminantium</i> M1	43	5	38	3
<i>Prevotella bryantii</i> B4	2	1	1	1
<i>Prevotella ruminicola</i> 23	6	1	5	1
<i>Ruminococcus albus</i> 7	12	4	8	1
Chromosome	7	2	5	1
pRUMAL01	4	2	2	0
pRUMAL02	1	0	1	0
pRUMAL03	0	0	0	0
pRUMAL04	0	0	0	0
<i>Ruminococcus albus</i> 8	15	6	9	4
<i>Ruminococcus flavefaciens</i> 007C	3	0	3	0
<i>Ruminococcus flavefaciens</i> 17	7	0	7	2
<i>Ruminococcus flavefaciens</i> FD-1	6	2	4	1
<i>Ruminococcus</i> sp. 18P13	9	3	6	1
<i>Wolinella succinogenes</i> DSM 1740	2	2	0	2
AC2a	7	2	5	0
ADA	6	1	5	1
AFa	5	0	5	0
AGa	15	4	11	0
AH	18	3	15	1
Ala	7	1	6	0
AJ	5	0	5	0
AMa	4	0	4	0
AN	4	0	4	0
APb	19	0	19	0
AQ	4	1	3	0
AS1a	7	1	6	0
ATa	4	0	4	0
AWa	2	0	2	0
BOa	17	1	16	0

a. CRISPR loci were detected by the CRISPRFinder online tool.

b. Cas proteins were detected by RAST.



Fig. 5. Putative CRISPR-associated (Cas) ORFs detected in the microbial reference genomes and the genome bins (Hoss et al., 2011).

© 2011 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology*, 14, 207–227

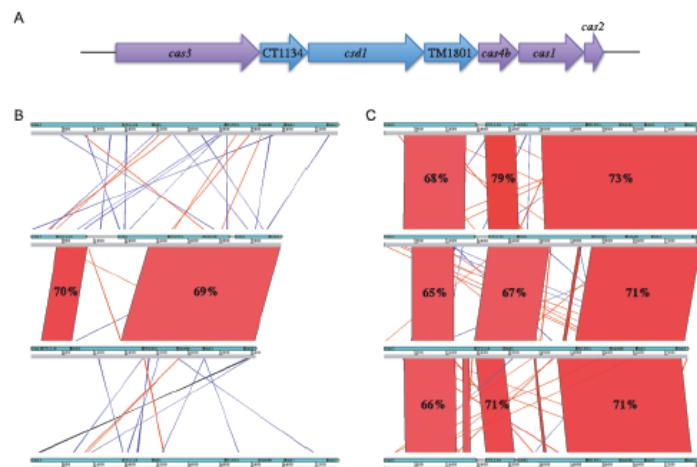
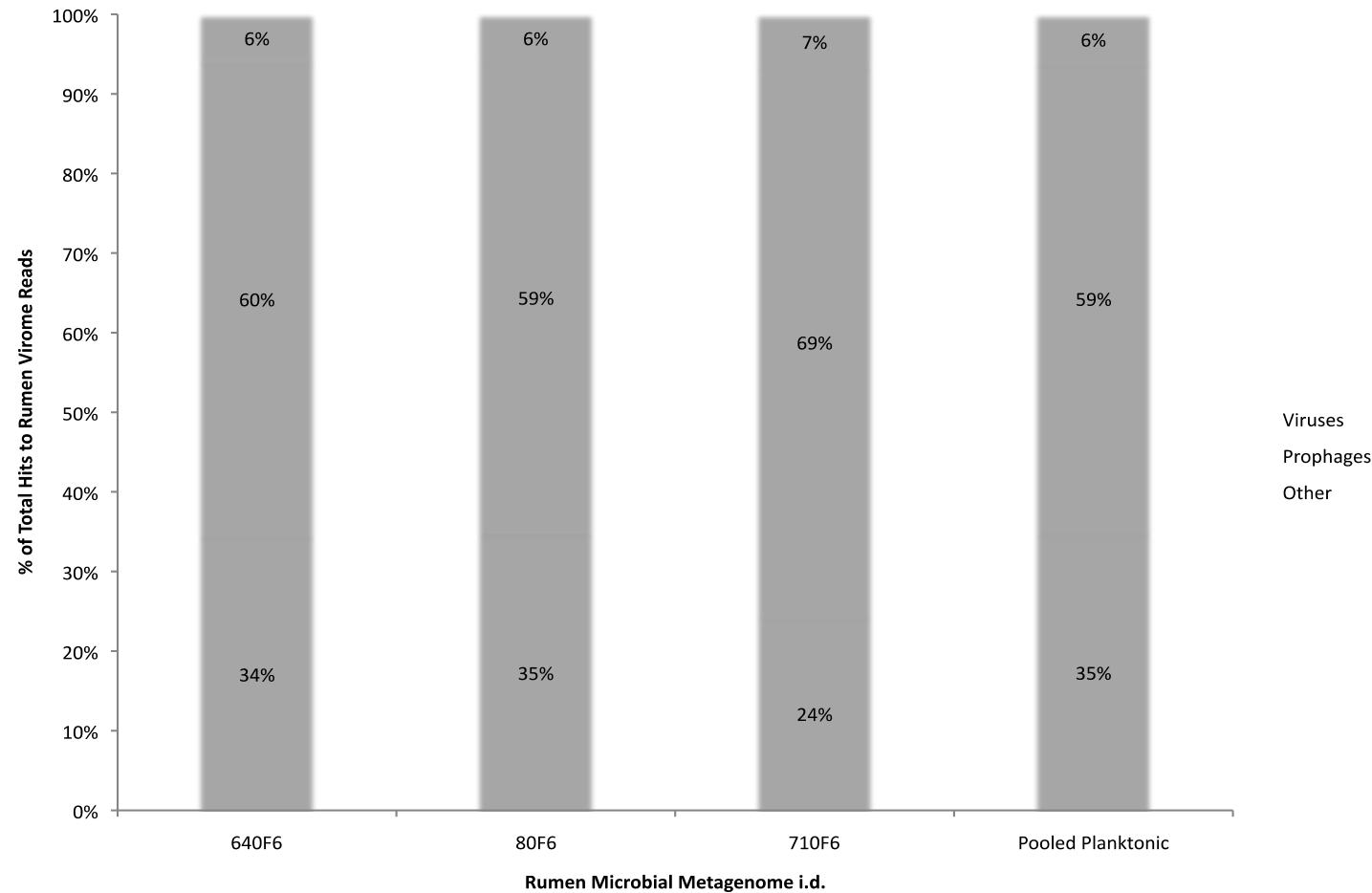
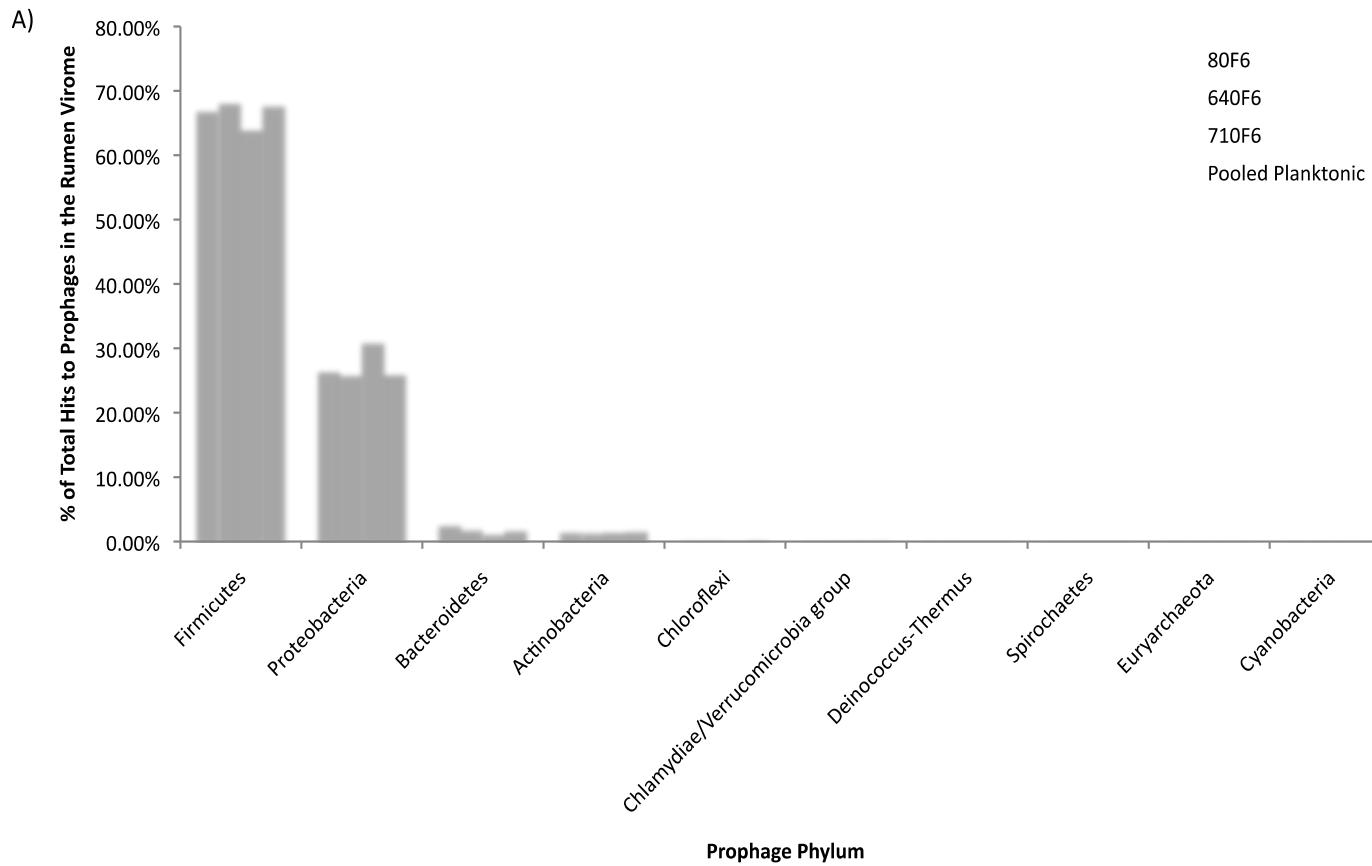


Fig. 6. CRISPR-associated (cas) gene organization in the genomes of *R. flavefaciens* strains FD-1 and 17, *R. albus* strains 7 and 8, and *Ruminococcus* sp. 18P13.  
A. The conserved arrangement of cas genes in the genomes of these five *Ruminococcus* strains.  
B. Comparisons of the sequence synteny were carried out using the Artemis Comparison Tool (ACT). The red and blue bands represent the forward and reverse matches between the nucleotide sequences respectively. The percentages denote the nucleotide sequence identity of that region between the two strains being compared. From top to bottom: *R. flavefaciens* FD-1, *R. flavefaciens* 17, *Ruminococcus* sp. 18P13, *R. flavefaciens* FD-1.

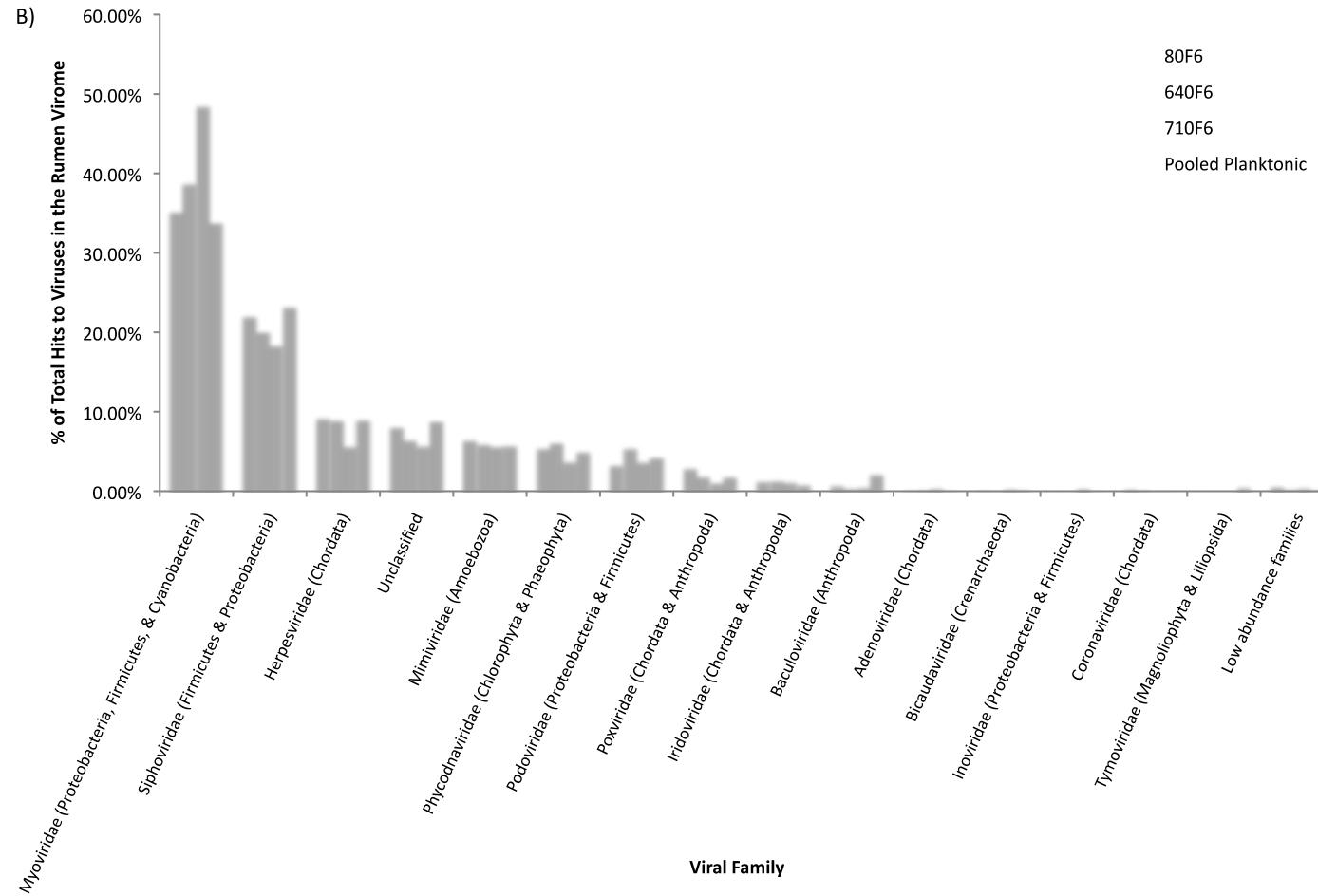
# Rumen Metagenomes vs. Virome



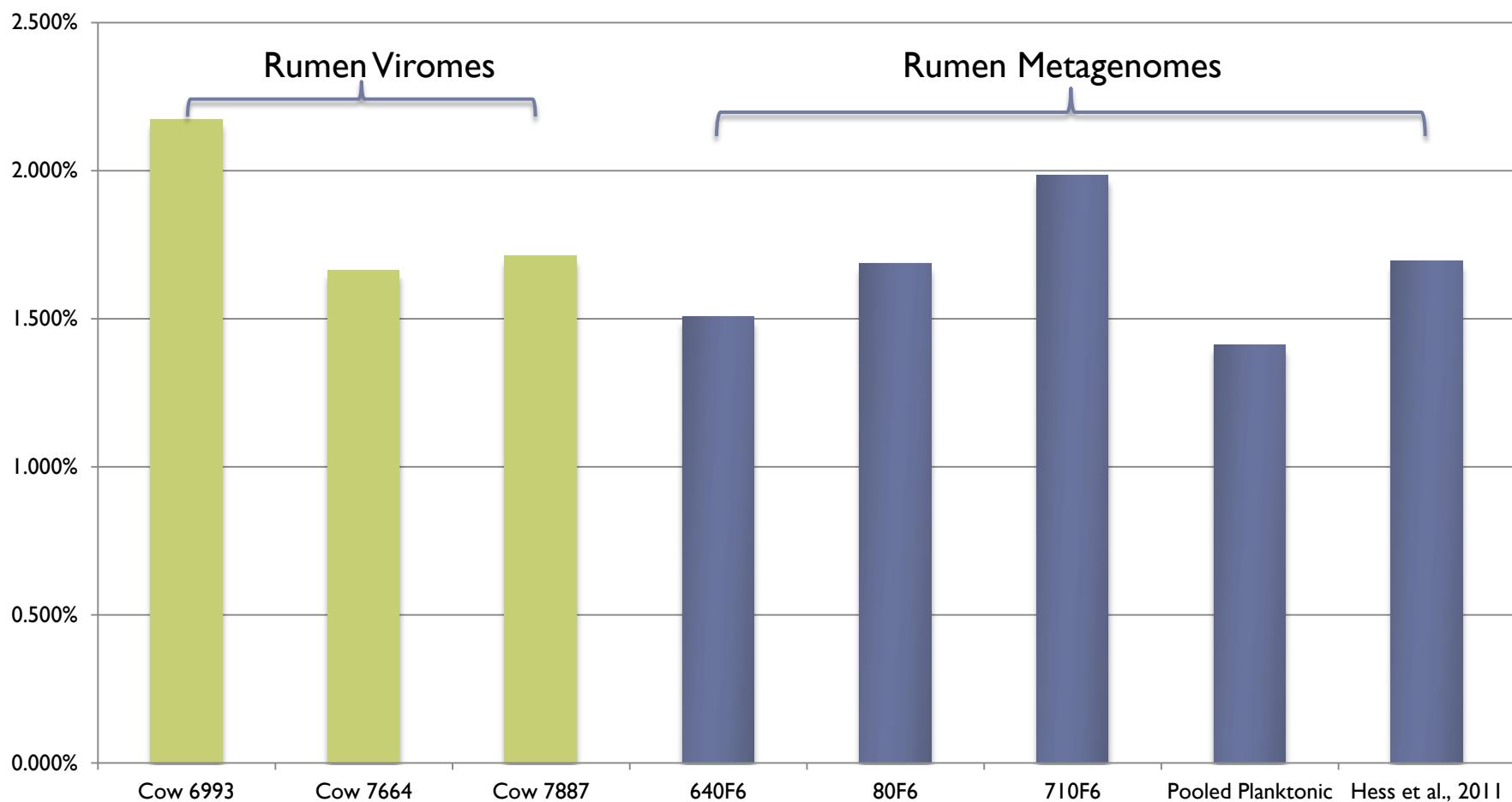
# Rumen Metagenomes vs. Virome - Prophages



# Rumen Metagenomes vs. Virome - Viruses

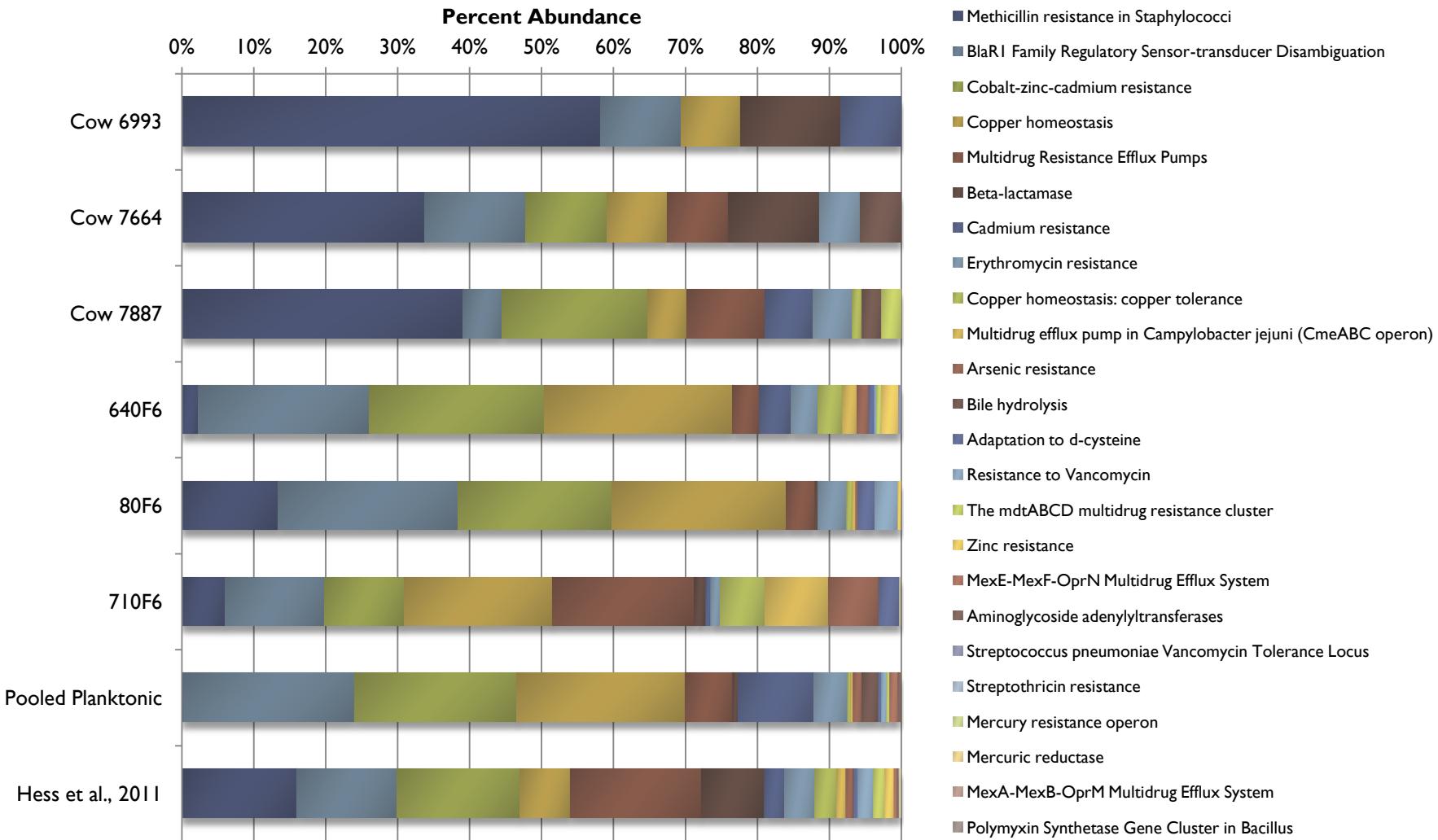


# Percent Abundance of Sequences in the Resistance to Antibiotics and Toxic Compounds Subsystem



Max. e-value cutoff = 1e-5

# Relative Abundance of Subsystems within the Resistance to Antibiotics and Toxic Compounds Subsystem



# Summary

---

- ▶ Rumen microbial and viral populations are closely associated
  - ▶ Changes in the structure of one community affects the other
  - ▶ Suggests “kill-the-winner” dynamic
- ▶ Evidence for phage-mediated HGT
  - ▶ Sequence similarities detected between microbes and viruses
  - ▶ Microbial sequences primarily displayed sequence similarities to prophages
  - ▶ Nearly all *Ruminococcus* mobile elements were similar to virome sequences



# Conclusions

---

- ▶ Rumen phage and microbial populations displayed a “kill-the-winner” dynamic
  - ▶ Presence of CRISPR in *Ruminococcus* strains
  - ▶ Dominant viral populations mirror the dominant microbial populations
- ▶ Evolution of the rumen microbiome is driven by phage-mediated horizontal gene transfer
  - ▶ Uncovered sequence similarities between rumen microbial genomes/metagenomes and the rumen virome
  - ▶ Majority of mobile elements in rumen genomes were similar to virome sequences
- ▶ Viruses represent a large unexplored area of genetic information that could be exploited for improving health and disease



# Acknowledgements

---

- ▶ Collaborators in Israel, Scotland, JCVI, JGI
- ▶ Funding: USDA-NRI, United States-Israel Binational Science Foundation, JGI's Laboratory Sequencing Program
- ▶ W. M. Keck Center – Alvaro Hernandez, Chris Wright
- ▶ Alejandro Reyes, Forest Rohwer – viral database

